

Study of the pre-processing impact in a Facial Recognition System

Guillermo Calvo¹, Bruno Baruque¹, and Emilio Corchado²

¹ Department of Civil Engineering, University of Burgos, Spain
gco0000@alu.ubu.es, bbaruque@ubu.es

² Department of Computer Sciences and Automatic, University of Salamanca, Spain
escorchado@usal.es

Abstract. The present work is a study of the influence of the preprocessing stage on the classification performance of a face recognition analysis. To carry out this task have made tests in a full FRS, evaluating each of its four stages and including several advanced alternatives in preprocessing, such as illumination normalization through the *Discrete Cosine Transformation* or alignment by *Enhanced Correlation Coefficient*, among others. The main goal of this work is determining how those different preprocessing alternatives interact with each other and in which degree they affect the overall Facial Recognition Systems (FRS). The tests make a special emphasis in using images that could have been obtained from a real environment, rather than at a lab environment, with the difficulties that this brings for facial recognition techniques.

Keywords: Face Recognition - Preprocessing - Normalization - Alignment - ECC - DCT

1 Introduction

A classification system is greatly influenced by data preprocessing and much of its success lies in selecting the best techniques for the task performed. As face recognition can be essentially put on the same level as a regular classification problem, it faces a similar challenge in this first stage. However, due to the nature of the images and the people involved, the data will inevitably include higher variations (or noise) than a classic classification problem. The images to analyse are influenced by other aspects intrinsic to the human physiognomy such as those due to the attitude of the model represented, changes in stands, gestures, clothing, hats, distance, tattoos, prosthetics or changes in appearance.

In addition, there are technical factors that increase the complexity of the classification. Many of them are due to image capture systems, associated systems lighting (flash) or kind of data from where images are obtained (still image, video, 3D, infrared, etc).

Finally, there are also external factors related to the environment such as the light conditions, background image, temperature or presence of other people, among others.

It is not always possible to work in a controlled environment so that images are not affected by the factors described above. However, not testing the system under these natural conditions will yield unreliable results.

There are several publications that conduct surveys on this kind of systems, but either they are centred in summarizing the results of the original publications [1] or they refer to specific problems observed in experimental tests [2,3]. Although these kind of studies are very interesting, it is also very informative to test the importance of the influence of all phases on the complete process of the final recognition of individuals. According to the knowledge of the authors, these kind of studies have not been carried out very often in a practical manner. In this case, the study is especially focused on the preprocessing of images, being this a crucial stage in the process.

2 Face Recognition Architecture

A Facial Recognition System is composed of five stages as it was discussed in [4]. These include the following ones (see also Fig. 1):

1. Image Capture
2. Facial Detection
 - Finding faces
 - Selection and image adjustment
3. Preprocessing
 - Illumination Normalization
 - Image Alignment
4. Feature Extraction³
5. Recognition process (classification)

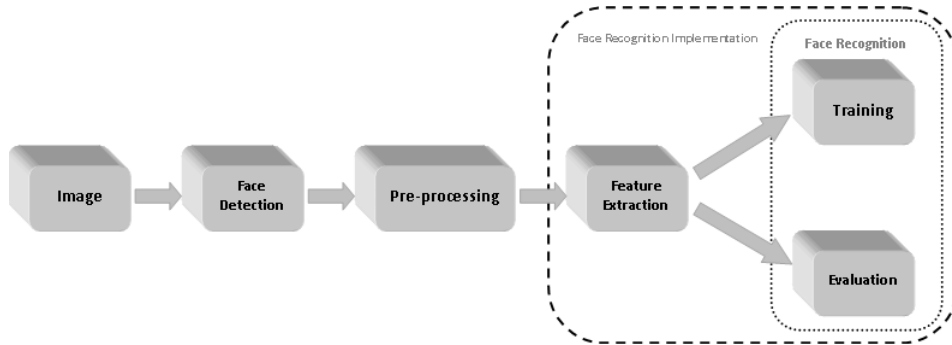


Fig. 1. Main scheme of a Face Recognition System

The capture of images rises the problems discussed in the introduction and a facial recognition process must be prepared to face them. The next step is to

³ It is usually performed with the recognition process stage.

discriminate the number of faces detected in the image. During this process the image is cropped to the size of the selected face and its size is standardized.

The resulting images must be preprocessed in order to obtain data to be easily classified. To do so, the light is homogenized to prevent glare or shadows areas that transform the image appearance. Another process within preprocessing is the alignment, which aim is assuring that the location of the face in a picture is the same in all pictures belonging the same class, in order to align facial features in all the images of the same person.

The stages of feature extraction and recognition are strongly interrelated and are usually included in one single step in practical approaches.

3 Techniques Used

3.1 Facial Detection

The *locally SMQT features & split up SNoW classifier* [5] has been used as the face detection technique in this study. This technique is based on appearance and is divided into two steps, the first is the use of SMQT [15] (Successive Mean Quantization Transform) locally in order to obtain the illumination invariant features and then, as second step, to use a SNoW (Sparse Network of Winnows) classifier.

- **SMQT**[15] This technique improves the image quality making it insensitive to the gain and the bias (off-set), considering an image that can be influenced by those factors (see eq. 1).

$$I(x) = gE(x)R(x) + b \quad (1)$$

The reflectance (R) has the the structure of the image itself, and it is needed to make the image invariable to gain (g) and bias (b) in order to be able to extract R and considering E as a *constant*.

- **SNoW Classifier**[5] This classifier get the features obtained by the previous step and uses a network of linear units to define the space of learned characteristics. To do so, it uses an initial SNoW classifier and the results from it are subdivided into other SNoW classifiers.

In the tests performed, this enables to crop only the face area detected, avoiding head shapes, hair, ears..etc.

3.2 Preprocessing

The purpose of this stage is to eliminate those features that hinder the classification of images of the same person (intra-class differences), thereby increasing the difference of them with others (inter-class differences).



Fig. 2. Image preprocessing secuency (*Initial*→ *Face Recognition*→ *DCT*→ *ECC*).

Illumination Normalization *Discrete cosine transform in logarithm domain* [6] along with the work of [7] [8] and a normalization method *based on RGB* [9] have been chosen as illumination normalization methods. The first one is based on a discretization of cosine as opposed to the second, which is based on a histogram with much milder changes in the values of illumination. Both methods perform a normalization of the full face image.

- **Discrete Cosine Transform (DCT)**[6] This normalization technique is based on the discrete Fourier transform, but using only real numbers. The procedure of this method is to adjust the dynamic range of a image in grayscale to the interval [0-255] and then truncate the ends of the image histogram⁴. This operation allows distributed gray levels along a image, eliminating the problems of existence of very bright values in the image that could dark the rest of the image after size changing. Finally, a photometric image normalization is done. The technique establish a predefined number of DCT coefficients to zero, eliminating the low frequency part of the data. This low frequency information is believed to be susceptible to changes in illumination.
- **RGB pixel compensation**[9] This method uses an adaptive illumination compensation, based on the black pixel, through a histogram equalization of the image. This is a two step method in which the first RGB image is compensated and then converted to YCbCr in order to normalize the image illumination.

Image Alignment Regarding the alignment techniques, *Enhanced Correlation Coefficient Maximization* [10] and *eye detection alignment* [11], which makes an alignment through the eye positions; have been chosen for this study. The first one was selected because is based on a template in contrast with the second one.

⁴ It removes a certain percentage of the low and high end of the histogram of an image.

- **Enhanced Correlation Coefficient (ECC)**[10]. The algorithm takes two images (an input image and an image template) as input, and estimates the 2D geometric transformation between them. It is possible to adjust several parameters in this algorithm, such as enabling implementation on a number of levels on a pyramid scheme or without it, the number of iterations per level, choose the type of transformation or using an initial transformation matrix.
- **Alignment through eye coordinates**[11]. The first step is to detect the eyes on a face using the cascade of Haar features. After obtaining those coordinates a transformation of the image is made to align it through a spatial transformation taking two checkpoints from the image.

3.3 Facial Recognition

Among the options for the recognition algorithm selected to construct the FRS for conducting this study, three different techniques have been included. Two holistic methods: *Eigenfaces*[12], as one of the most widespread algorithms on this category, and *Fisherfaces*[13] which is a well-known evolution from the previous one. And, to broaden the scope of the study, a feature based method such as *Hidden Markov Model*[14] has also been selected.

- **Eigenfaces**[12] (Holistic) It is a classification method based on Principal Component Analysis (PCA) to reduce the dimensionality⁵ of each image and projected their attributes on the new dimensions considered. Finally, the final classification is obtained by comparing the Euclidean distances between the data obtained for each image. This technique provides a reasonably satisfactory results [12] and has low computational load.
- **Fisherfaces**[13] (Holistic) The Fisherfaces method is based on LDA (Linear Discriminant Analysis) and it uses information between members of the same class to develop a set of feature vectors where the variations between the different faces (or classes) are emphasized while the differences within the same class are minimized. Previously, Fisherfaces uses PCA to reduce the dimension of the data, as can be seen on Eq. 2, where S_B is the scattering matrix between classes, S_T is the intra-class ones and W is the orthonormal matrix of the new space. The results are better than those achieved just with PCA, without preprocessing as varying lighting conditions or with slight changes in facial expressions [13].

$$W_{opt} = \arg \max \frac{|W^T S_B W|}{|W^T S_W W|} \quad (2)$$

- **Hidden Markov Model (HMM)**[14](Feature Based). A Hidden Markov Model, is a statistical model which assumes that the system model is a Markov process of unknown parameters and could be considered as a dynamic Bayesian simple network. This algorithm is based on the division of

⁵ Having selected 50 training images the number of principal components used in this algorithm is 49.

7 facial image regions (7 nodes) and a state transition probability. The algorithm provides a probability that a given region of the face will follow another with certain determined features. According to that probability, the analysed image is associated or not to the considered matching image. This algorithm previously uses a histogram equalization (HE).

4 Experiments and Results

To establish a valid test methodology, given that the problem to solve is a multiclass classification one, a comparison of One-Vs-All (OVA) is performed, as suggested in [16]. In all tests performed, a cross-validation 6 K-fold [17] for 10 different people (or classes) has been performed. The clusters are composed by a set of 60 images (6 per person), using 50 for training and 10 for testing. The confidence interval values are represented by percentage of success $(1 - E) \cdot 100$ and using the mean error obtained through $E = \frac{1}{K} \sum_{i=1}^k E_i$, where E is the error percentage and k the number of folds in the cross-validation.

4.1 Databases

Two different image databases have been used for testing. The first is the **Caltech** [18] one which shows people with different image backgrounds, light conditions, facial expressions and camera distances. These images are not preprocessed, maintaining consistency with those we could take in any environment, which increases the difficulty for a FRS discriminating among them ⁶. The second database used on this work is **ORL** [19], where pictures have all the same image background, are focussed at heads and, although they are taken at the same distance, the pose variation is much bigger than those used in *Caltech*.

4.2 Experiments

Two experiments have been performed for this study. The first one has been composed as the most complete test possible, using all combinations available to construct the FRS. Results obtained in this experiment have been confirmed in the second one, with the use of a different database and choosing the combinations with the best results obtained in the first experiment.

Procedure Experiment 1. The images have been transformed to grayscale when it was necessary, also they have been resized to 179x118 pixels in its initial stage and to 46x46 pixels from the face detection process in order to reduce the computational load.

The objective of this experiment is to observe how the recognizing accuracy of the system increases or decreases for each of the classifiers selected (see 3.3)

⁶ The high intra-class variation increases the difficulty of classification methods that rely on common features in each class (Fisherfaces).

when additional stages are added to the FRS. In order to observe this variations, the following test configurations have been included in the test:

- Original images only (*INITIAL*)
- Cropped facial images -only face- (*FACE DETECT.*)
- DCT illumination normalization algorithm (*DCT*)
- RGB illumination normalization algorithm (*NORM.RGB*)
- ECC alignment system (*ECC*)
- Eye Position alignment system (*EYE.ALIGN*)
- DCT illumination correction + ECC alingment (*DCT + ECC*)
- DCT illumination correction + Eye-align (*DCT + EYE.ALIGN*)
- Norm-RGB illumination correction + ECC alingment (*NORM.RGB + ECC*)
- Norm-RGB illumination correction + Eye-align (*NORM.RGB + EYE.ALIGN*)

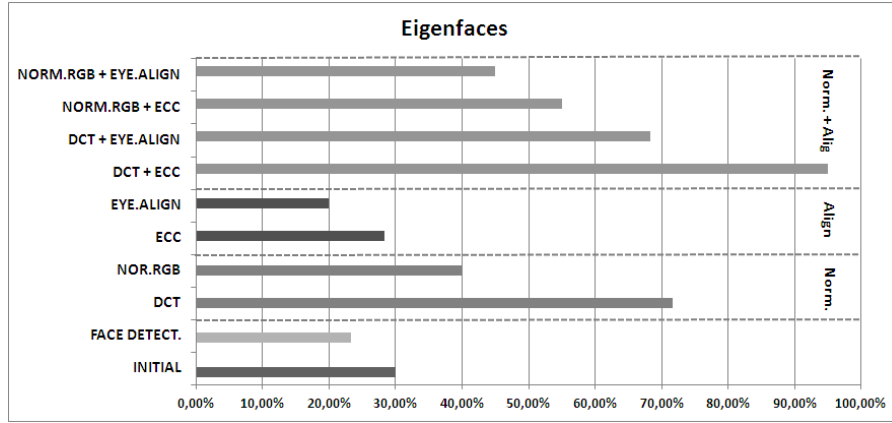


Fig. 3. Eigenfaces Results

Results Experiment 1. The results obtained in this test are shown in Figs. 3, 4 and 5. In them, the corresponding confidence rates are shown.

Comparing results from the initial stage to those including face detection ones show that:

- The results for eigenfaces (see fig. 3) are worse to others because in all processes the outline of the head, which should help this particular classifier to discriminate people, has been removed from the pictures.
- For fisherfaces (see fig. 4) the results are greatly improved (from 8.33% to 46.66%) because the gain of minimizing similar data inside each class due to same image backgrounds is bigger than the loss due to the absence of head shape.
- In the HMM model (see fig. 5) its results are also highly improved (from 18.33% to 63.33%), because this classification technique is based on features and the elimination of the image background increases their discrimination based on probabilities.

Comparing the results from the combinations using only the face detection step to those including illumination normalization, shows that:

- For eigenfaces (see fig. 3) there is a high increase in the success rate, which goes from 23,33% to 71,66% for the best technic selected (DCT), due to the improving of data (avoiding hidden areas) making them linearly divisible.
- In fisherfaces (see fig. 4) the rate of success decrease because some distinguishing features between classes are erased by illumination normalization process.
- For HMM (see fig. 5) there is a slight improvement in the success rate (increment 10%), though this happens just with the best method (DCT). This is due to improved of using both DCT and the pre-filter algorithm itself (see 3.3).

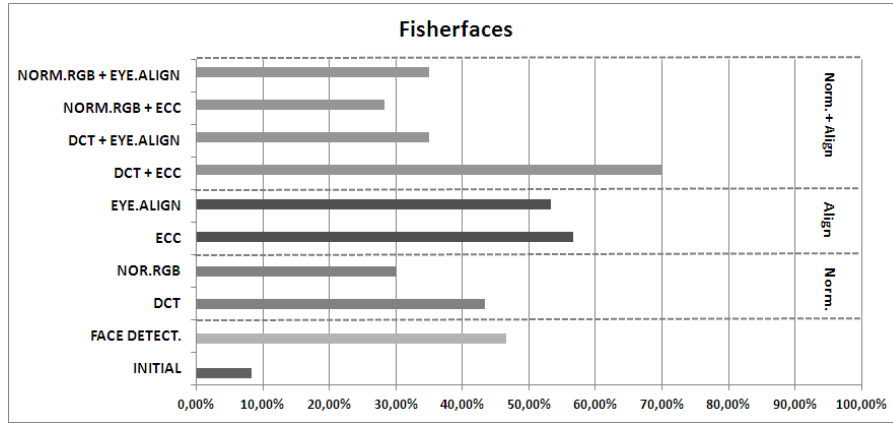


Fig. 4. Fisherfaces Results

In the results for combinations using the normalization step without a previous illumination normalization show that:

- For eigenfaces (see fig. 3) shows just very slight improvement rates or even some worse with the technique (EYE.ALIGN). That indicates without a previous illumination normalization the alignment does not increases the separation between classes.
- In fisherfaces (see fig. 4) the image alignment stage, however, no worse affects and even get better results in both techniques (ECC y EYE.ALIGN) going from 46,66% to 56,66%.
- HMM (see fig. 5) does not show improvement in its performance and even there is a slight decrease for the worst method (EYE.ALIGN) by the introduction of black pixels zones inside images in this step, this is something which makes difficult to HMM to discriminate between classes.

The last combination set, including a illumination normalization and then an alignment method, show that:

- Eigenfaces (see fig. 3) shows better results than any of the other combinations for (DCT + ECC). The alignment prior and subsequent illumination normalization makes first increase the separation between classes and then decrease the distance within each class.
- In fisherfaces (see fig. 4) the combination of illumination normalization and then applying alignment makes the results improve to 70% for (DCT + ECC).
- For HMM (see fig. 5) there are not significant differences in the values obtained and show a substantial decrease for the worst combination (NORM.RGB + EYE.ALIGN).

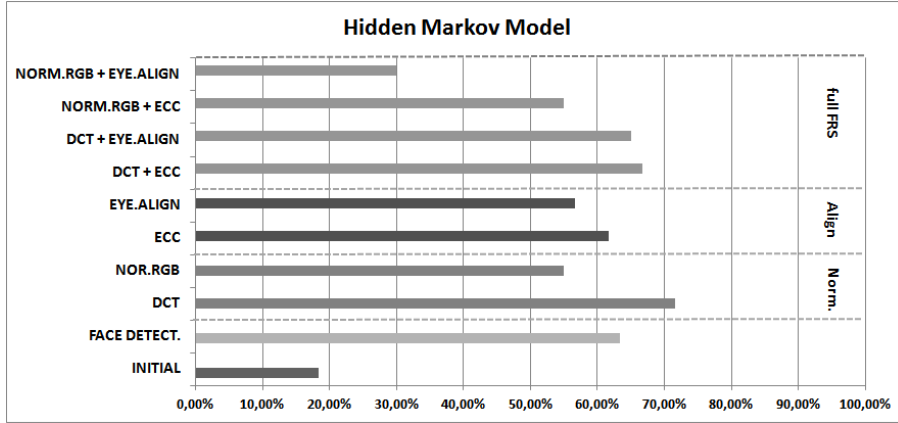


Fig. 5. Hidden Markov Model Results

It is obvious when analysing these results that the preprocessing stage greatly affects final results for eigenfaces classifiers and that the biggest impact within it is the illumination normalization.

It can also be observed that preprocessing decisively affects less fisherfaces than eigenfaces, since images are standardized and make fisherfaces lose its more defining quality: to increase inter-class differences (which are now much less evident), obtaining worse results when explicit different position variation is included within a single class (high intra-class variability in initial images).

It can also be noted that the results for the HMM method are not strongly influenced by preprocessing stages and that the algorithm is able to get acceptable results without the use of this stage, since it is based on features, and analyzing them separately is not severely affected by large areas of lighting changes or alignment.

Procedure Experiment 2. In this experiment, images are initially in grayscale and resized to 46x46 pixels from the face detection process.

For this test all mentioned algorithms from the stage of face detection to the illumination normalization and alignment have again been considered. This time, only the best combination of each stage has been used (*Face Recognition* →

$DCT \rightarrow ECC \rightarrow DCT+ECC$). The purpose of this second study is to corroborate the results obtained in the *experiment 1* (see Section 4.2).

	Face-Detect.	Norm-DCT	Align-ECC	DCT+ECC
Eigenfaces	60%	70%	71%	93%
Fisherfaces	55%	31%	65%	60%
Hidden Markov Model	85%	60%	68%	67%

Table 1. Comparison for the best solution with ORL images.

Results Experiment 2. As seen from the data given in Table 1, classification methods which are inherently better avoiding *noise* in lighting, position or gestures (fisherfaces and HMM) get worse results than eigenfaces after preprocessing. This emphasizes the importance of preprocessing methods used when the classifier is not oriented on differences between classes or features. These results are consistent with those in the previous experiment and reflect equally the importance of preprocessing for FRS.

5 Conclusions

It follows from the results obtained that some preprocessing methods do not work properly with the classifiers that take into account the knowledge of the number of classes (Holistic) or are not so critically dependent on it for their operation (Features Based). However, methods such as eigenfaces crucially improve their classification capabilities with a suitable preprocessing.

Feature based or holistic methods designed to avoid intrinsically variations introduced by lighting or pose, are most robust in results without a prior preprocessing, being these more suitable than methods that have not been developed with this purpose (such as eigenfaces).

With the existence of many variations in image conditions for the same class (which implies a great intra-class variability) may be more beneficial to use a preprocessed and holistic classifier than using a discriminatory or advanced classifier (such as fisherfaces or HMM), because the preprocessing makes the classes are linearly separable without leaving that task to the last stage of the FRS.

When the conditions of image capture do not include a controlled environment, classification methods find serious problems in getting good levels of confidence that maybe a good preprocessing could fix it.

The preprocessing stage is not just important, but also delicate, it is essential to know the functioning of the classifier used to choose an appropriate preprocessing to improve the results, in case of mistake its impact could be negative.

Acknowledges

This research is partially supported by the Spanish Ministry of Economy and Competitiveness under project TIN2010-21272- C02-01 (funded by the European

Regional Development Fund), SA405A12-2 from Junta de Castilla y León. This work was also supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070).

References

1. Zhao, W.; Chellappa, R.; Phillips, P. J. and Rosenfeld, A. Face recognition: A literature survey. *ACM Computing Surveys*, 35, 399-459 (2003)
2. Zhang, X. and Gao, Y. Face recognition across pose: A review. *Pattern Recognition*, 42, 2876 - 2896 (2009)
3. Sang-Il, C.; Chong-Ho, C. and Nojun, K. Face recognition based on 2D images under illumination and pose variations. *Pattern Recognition Letters*, 32, 561 - 571 (2011)
4. Gonzalez, Rafael C. and Woods, Richard E., *Digital Image Processing*. Addison-Wesley Pub (Sd) (2008)
5. Nilsson, Mikael and Nordberg, Jorgen and Claesson, Ingvar, *Face Detection using Local SMQT Features and Split up Snow Classifier* (2007)
6. Chen, W. and Er, Meng J. and Wu, Shiqian, *Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain*. *Systems, Man and Cybernetics, Part B, IEEE Transactions on* (2006)
7. V. Štruc and N. Pavešić, *Photometric normalization techniques for illumination invariance*. *Advances in Face Image Analysis: Techniques and Technologies, IGI-Global* (2011)
8. V. Štruc and N. Pavešić, *Gabor-based kernel-partial-least-squares discrimination features for face recognition*. *Informatica (Vilnius)* 2009
9. Chen, Lihui and Grecos, Christos, *Fast skin color detector for face extraction*. *Electronic Imaging* (2005)
10. Evangelidis, G.D. and Psarakis, E.Z., *Parametric Image Alignment Using Enhanced Correlation Coefficient Maximization*. *Systems, Pattern Analysis and Machine Intelligence, IEEE Transactions on* (2008)
11. Everingham, M. and Sivic, J. and Zisserman, A., “Hello! My name is... Buffy” – Automatic Naming of Characters in TV Video. *Proceedings of the British Machine Vision Conference* (2006)
12. Turk, M. and Pentland, A., *Eigenfaces for Recognition*. *Journal of Cognitive Neuroscience* (1991)
13. Belhumeur, Peter N. and Hespanha, João P. and Kriegman, David J., *Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection* (1997)
14. H. Miar-Naimi, and P. Davari, and AWT-TAG, *A New Fast and Efficient HMM-Based Face Recognition System Using a 7-State HMM Along With SVD Coefficients*. *Iranian Journal of Electrical & Electronic Engineering* (2008)
15. Nilsson, M. and Dahl, M. and Claesson, I., *The successive mean quantization transform*. *Acoustics, Speech, and Signal Processing. Proceedings (ICASSP '05). IEEE International Conference* (2005)
16. Rifkin, Ryan and Klautau, Aldebaro, *In Defense of One-Vs-All Classification*. *Journal of Machine Learning Research* (2004)
17. Refaeilzadeh, Payam and Tang, Lei and Liu, Huan, *Cross-Validation*, *Encyclopedia of Database Systems*. Springer US (2009)
18. Caltech Computational Vision Group. *Faces 1999 Database*. <http://www.vision.caltech.edu/html-files/archive.html> Last accessed: 2012
19. AT&T Laboratories. *Cambridge ORL Faces Database*. <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html> Last accessed: 2012